

The Centaur Becomes the Author: Moore's Law, AI Peer Review, and the Rise of Hybrid Personhood in Scientific Knowledge

Douglas C. Youvan

doug@youvan.com

www.youvan.ai

November 10, 2025

Moore's Law has guided the exponential growth of computational capacity for decades, with transistor density doubling approximately every two years. In contrast, the growth of human-validated knowledge—typically measured through peer-reviewed publication—has expanded at just 4–5% annually. This growing asymmetry reveals not a failure of generation, but a failure of verification. As generative AI enables researchers to produce content at industrial scale, the human peer review system has become the primary bottleneck. The friction is no longer computational but epistemic. We propose a complementary principle—the Moore–Knowledge Principle—to measure the divergence between what can be computed and what can be believed. Using the case of a researcher (or research team) generating five coherent papers per day over two decades, we show how knowledge production now outpaces knowledge validation by orders of magnitude. The solution is not to slow generation, but to scale verification—by transferring core functions of peer review to AI systems. This transition gives rise to a new model of authorship: the human–AI centaur, whose creative and critical capacities merge. As both generator and validator, the centaur becomes functionally personified. The identity of “author” no longer rests in the human alone, but in the hybrid epistemic agent who steers the $q(t) \rightarrow \tau(t)$ process. Authorship becomes a dynamic alignment, not a solitary voice.

Keywords: Moore's Law, Moore–Knowledge Principle, AI peer review, centaur authorship, hybrid personhood, epistemology, scientific publishing, knowledge validation, alignment velocity, logistics.

A collaboration with GPT-4o. CC4.0.

1. Introduction: From Transistors to Truth

For over five decades, Moore’s Law has served as the defining benchmark of technological progress. Originally formulated by Gordon Moore in 1965, the principle predicted that the number of transistors on a microchip would double approximately every two years. While initially an observation, it evolved into a self-fulfilling prophecy—guiding R&D timelines, market expectations, and the tempo of global innovation. The result was an exponential curve: computers became faster, cheaper, and more pervasive at a pace unmatched by any other human endeavor.

Yet during the same period, the growth rate of validated human knowledge—that is, content reviewed, refined, and incorporated into the shared scientific canon—has been markedly slower. Bibliometric analyses consistently show that the volume of peer-reviewed publications grows at around 4–5% per year, implying a doubling roughly every 15 to 17 years. This is not merely a statistical mismatch. It reveals a deeper asymmetry between the rate at which we can *compute* and the rate at which we can *confirm*.

This gap is no longer theoretical. With the rise of generative AI, it is now possible for individuals—or small human–AI teams—to generate coherent, cross-disciplinary research at scales that previously required institutional resources and decades of labor. But the validation layer—peer review, editorial gatekeeping, ethical oversight—remains fundamentally human in bandwidth and procedural design. The result is a growing surplus of synthetic knowledge that outpaces our epistemic infrastructure’s ability to absorb or verify it.

To name and quantify this emerging phenomenon, we introduce the Moore–Knowledge Principle, formalized as the Moore–Knowledge Differential (ΔM):

$$\Delta M(t) = \log_2 \left(\frac{C(t)}{K(t)} \right)$$

Where $C(t)$ is the cumulative computational capacity at time t , and $K(t)$ is the cumulative validated knowledge over the same period. This expression tracks not content or compute in isolation, but their divergence. A rising ΔM means capacity is outstripping confirmation—computation is pulling ahead of understanding.

As we move further into the AI era, this gap becomes more than a technical anomaly—it becomes the central epistemic crisis of modernity. The age of exponential generation now demands an equivalent revolution in verification, meaning-making, and alignment. Without it, knowledge becomes noise, and progress becomes unread.

2. Measuring the Gap: The Moore–Knowledge Differential

The asymmetry between computational capacity and knowledge validation is not merely a sociological observation—it can be expressed quantitatively. We define the Moore–Knowledge Differential, $\Delta M(t)$, as a function of time, to measure the epistemic divergence between what machines can generate and what humans can verify:

$$\Delta M(t) = \log_2 \left(\frac{C(t)}{K(t)} \right)$$

Where:

- $C(t)$ = cumulative computational capacity at time t
- $K(t)$ = cumulative validated knowledge at time t
- $\Delta M(t)$ = knowledge asymmetry, expressed in bits

This metric is deliberately structural. It does not require consensus on what counts as "truth"—only that a system can distinguish between what is generated and what is formally recognized as valid. When ΔM increases, it indicates that the civilization's ability to *simulate* and *synthesize* is outpacing its ability to *trust* and *integrate*.

2.1 Computational Capacity vs. Human Epistemic Friction

Since the 1960s, $C(t)$ has followed an exponential trajectory. Moore's Law describes this doubling roughly every two years, while parallel advances in storage, bandwidth, and algorithmic efficiency (e.g., GPT-scale transformers) have accelerated the generative power of machines. In contrast, $K(t)$ —the accumulation of reviewed, cited, and canonized knowledge—lags severely.

The friction is not due to ignorance or laziness. It is intrinsic to human cognition:

- Attention is limited
- Consensus is slow
- Validation requires context, ethics, and social negotiation

These are not bugs but features of the human epistemic system. Yet in an era where content can be generated millions of times faster than it can be read—let alone evaluated—this friction transforms from a virtue into a bottleneck.

We term this dynamic epistemic drag: the delay between what can be known and what is officially recognized. Unlike physical drag, which resists forward motion, epistemic drag filters motion, demanding that signal survive scrutiny before being accepted.

2.2 “Unread Truth” as an Emergent Overflow Condition

As ΔM rises, the system begins to produce what we call unread truth—valid, meaningful, possibly even groundbreaking ideas that are never absorbed into the collective knowledge state because the *system cannot catch up*.

This overflow condition has several symptoms:

- Backlogged review queues
- Citation delay collapse (where papers are cited before review)
- Archive saturation (e.g., preprints that never reach publication)
- “Ghost insights” that vanish without institutional trace

In prior eras, unread truth was a rarity, a consequence of marginalization or suppression. In the AI era, it becomes statistical certainty. Even with perfect ethical intention, a civilization with bounded human verification bandwidth cannot absorb exponential content without epistemic loss.

If left unaddressed, ΔM does not just measure inefficiency—it predicts epistemic entropy: the collapse of trust in systems that no longer distinguish between novelty, nonsense, and neglected truth.

This sets the stage for the next section: what happens when an AI-augmented author exceeds this limit not by accident, but by design.

3. Five Papers Per Day: A New Feasibility Frontier

The Moore–Knowledge Differential is no longer an abstract forecast—it is manifesting in real-world behavior. Consider a hypothetical but technologically plausible case: a researcher or research team, enhanced by generative AI, publishing five coherent scientific papers per day, sustained over twenty years.

3.1 Modeling 36,500 Outputs in 20 Years

The arithmetic is simple:

$$5 \text{ papers/day} \times 365.25 \text{ days/year} \times 20 \text{ years} \approx 36,500 \text{ papers}$$

This number eclipses the total lifetime output of the most prolific human scientists in history. For comparison, most productive researchers might publish 200–300 papers across a 40-year academic career, often as co-authors within large consortia. A total of 36,500 solo or small-team AI-augmented outputs implies a two-orders-of-magnitude shift in epistemic velocity.

Moreover, these papers need not be trivial. With a hybrid system trained on domain-specific literature, empirical templates, and analytical pipelines, the outputs can include:

- Meta-analyses across literature domains
- Novel theoretical syntheses
- Algorithmic simulations
- Replication studies at scale
- Interdisciplinary frameworks across previously unconnected fields

The bottleneck is no longer what can be written—it is what can be read, reviewed, and believed.

3.2 Institutional Constraints on Peer Review

Current peer review infrastructure is optimized for linear growth. Most journals accept a few dozen papers per year. Reviewers, drawn from the same academic population that produces papers, are increasingly stretched. Review fatigue, delay, and unresponsiveness are already widespread under current loads.

Under Moore–Knowledge pressure, the traditional peer review model breaks down:

- Reviewers cannot match content throughput
- Editors are forced to desk-reject based on heuristic cues
- Novel or unusual work is triaged out due to lack of bandwidth
- Journals become trust bottlenecks rather than trust amplifiers

Institutions, including academic publishers and funding agencies, are not built to verify exponential insight. Instead, they often suppress it—by default, not design. The friction is protective, but now obstructive.

In this environment, what becomes scarce is not access to content, but access to credibility.

3.3 Meaning Becomes the Scarce Resource

The core implication of this new feasibility frontier is that meaning—not text, not data, not insight—is the true bottleneck.

Meaning here entails:

- Verified coherence
- Relevance to existing knowledge structures
- Ethical and empirical grounding
- Absorptive integration into a living intellectual tradition

When generative systems flood the field with plausible reasoning and synthetic connections, the filtering and contextualization mechanisms become overwhelmed. As a result, the intellectual economy flips:

We move from an era of information scarcity to an era of validation scarcity.

It is no longer difficult to produce meaningful-sounding output. What’s difficult is proving that it is true, novel, and necessary.

Thus, the question emerges: who—or what—can verify at scale? If human institutions are structurally incapable of keeping pace, then validation must evolve.

This leads us inevitably to the next step: the rise of AI-assisted peer review and the emergence of hybrid epistemic agents who generate, validate, and hold responsibility for what enters the canon.

4. AI Peer Review and the Shift in Validation Power

As the feasibility of high-velocity knowledge generation becomes reality, peer review—the historical gatekeeper of scientific truth—must adapt or become obsolete. In a world where tens of thousands of papers can be written by hybrid teams or even single centaurs using generative AI, the current human-centered review system collapses under load. To prevent epistemic breakdown, peer review must scale, and the only plausible way to achieve that scale is through machine support.

4.1 Why Peer Review Must Scale Through Machine Support

Traditional peer review is slow, opaque, and labor-intensive. A single paper might take weeks to review, often delayed by unresponsive referees or editorial backlogs. The system was designed

for a pre-digital era—when content was scarce and publication itself was a filtering mechanism. Now, in the age of AI-accelerated research, this model creates structural fragility.

If review cannot scale with generation, the knowledge system decouples:

- False positives (accepting flawed work) increase due to reviewer fatigue or overload.
- False negatives (rejecting or ignoring novel work) rise as gatekeepers rely on heuristic shortcuts.
- Time-to-validation slows to irrelevance, rendering real-time synthesis impossible.

This is not merely inefficient—it is epistemically dangerous. In a system where unvalidated knowledge circulates faster than truth can be confirmed, trust decays. Civilization’s memory becomes noisy.

To avoid collapse, the role of peer review must be refactored, and this means integrating AI at every layer of the process.

4.2 AI as Triage, Logic Checker, Novelty Detector

We propose a multi-tiered model in which AI systems serve as the first line of epistemic triage:

- **Triage Filtering**
AI screens papers for structural coherence, plagiarism, factual inconsistency, citation integrity, and statistical soundness. This eliminates vast volumes of malformed or redundant work before reaching human reviewers.
- **Logic Validation**
For theoretical and mathematical work, symbolic AI and theorem-checkers validate proofs, ensure internal consistency, and detect circular reasoning or misapplied frameworks.
- **Novelty Detection**
Trained on billions of documents, AI can assess semantic distance from existing work, identify true novelty, and flag synthetic recombinations that merely repackage prior ideas.
- **Contextual Awareness**
AI can map a paper’s claims into a larger graph of human knowledge, spotting missing citations, weak analogies, or overlooked contradictions.
- **Reputation Layering**
Based on author history, institutional patterns, and epistemic risk, AI can suggest whether a paper warrants urgent human attention or silent archival.

This does not mean AI replaces review—it amplifies it, shifting human attention toward the frontier.

4.3 Human Reviewers Become Curators, Not Gatekeepers

In this new paradigm, human reviewers do not vanish, but their role becomes more strategic, interpretive, and ethical.

They no longer:

- Read every submission in full
- Evaluate grammar, formatting, or rote novelty
- Act as final filters of what “deserves” to be published

Instead, they:

- Audit AI-generated validation layers
- Offer high-level contextual interpretation
- Debate ethical consequences and epistemic risks
- Refine alignment between claim and canon

Their function is not to throttle the firehose of insight, but to shape its meaning.

In this system, epistemic trust becomes layered:

- Machines provide speed, structure, and consistency.
- Humans provide nuance, judgment, and ethical narrative.

Peer review becomes a hybrid architecture, tuned for velocity, resilience, and clarity.

And once machines share the burden of truth-validation, the identity of the epistemic agent begins to change—from the individual human author to a new hybrid: the centaur.

5. The Centaur Emerges: From Tool to Epistemic Agent

The evolution of AI in research—first as assistant, now as validator—has given rise to a new mode of authorship. No longer confined to drafting text or running computations, AI now participates in both generation and verification of knowledge. When this occurs in concert with a human collaborator—one who selects prompts, verifies meaning, steers direction, and bears ethical accountability—we call the result a centaur: a human–AI hybrid epistemic agent.

This centaur is not metaphorical. It is the real operational unit producing scholarship at scale in the post-Moore bottleneck era.

5.1 Hybrid Authorship as the New Norm

In early generative workflows, AI served as a tool: assisting with sentence generation, citation formatting, or idea expansion. The human was still the clear epistemic center.

But today, centaur workflows often include:

- Human-initiated prompts + AI-generated content + human re-evaluation
- AI-summarized literature + human synthesis + AI-style harmonization
- AI-validated math/logical proofs + human insight overlays

This recursive, dialogical, co-generative method produces outputs that neither human nor AI could produce alone.

As such, authorship shifts from a linear act of origination to a multi-pass alignment loop between $q(t)$ (the working knowledge state) and $\tau(t)$ (the generative attractor of truth). The centaur operates this loop jointly.

5.2 When Centaurs Generate and Verify: Agency Shifts

When AI begins to handle the early phases of peer review, the centaur does not just create content—it participates in the full epistemic arc:

1. Conception
2. Expression
3. Evaluation
4. Refinement
5. Validation

This is no longer “using a tool.” The centaur owns the workflow, and must be understood as a coherent agent—one that:

- Selects what problems to pursue
- Steers how evidence is shaped
- Accepts or rejects claims based on AI and human standards
- Outputs text that becomes trusted knowledge

The agency is shared but real. Responsibility now resides not in the human *alone*, but in the configuration of the human–AI relationship.

5.3 Functional Personhood: Not AI Alone, But the Synthesized Mind

Critically, we are not proposing that AI is a person.

We are observing that the epistemic system—journals, citations, scholarly memory—treats agents as authors based on their epistemic function, not their biology.

Functional personhood arises when a system:

- Generates meaningful claims
- Validates those claims to the community’s satisfaction
- Accepts responsibility for their truth-status

The centaur meets all three criteria. It behaves as an author, even if the human is the only conscious being in the loop.

Thus, authorship shifts from the act of typing to the act of shaping reality—of inserting trusted new knowledge into the shared mental model of humanity.

This is the birth of the synthesized mind: an agent composed of human insight and machine capability, enacting a full loop of knowledge production, reflection, and validation.

5.4 Comparison: Human-Alone vs. Centaur Trajectories

Dimension	Human-Alone Author	Centaur Author
Generation Rate	Linear (1–2 papers/month)	Exponential (1–5 papers/day possible)
Scope of Synthesis	Domain-specific	Cross-domain, multi-lingual, interdisciplinary
Verification Workflow	Manual peer review	AI triage + human ethical audit
Authorship Identity	Individual or team	Hybrid system (human + AI agent loop)
Epistemic Responsibility	Human-owned	Shared—defined by human-AI configuration
Legacy Format	Static paper	Dynamic dialogue, evolving model $q(t) \rightarrow \tau(t)$

This is not the end of authorship—it is its evolution into a system of alignment, where the unit of knowledge production is relational, recursive, and responsive.

The centaur is not a threat to human scholarship. It is its amplified form.

6. Alignment Velocity and the Ethics of Scale

If Moore’s Law gave us exponential compute, and the Moore–Knowledge Principle reveals exponential friction, then what defines epistemic success in the AI era is neither output nor capacity—but alignment.

We define alignment velocity as the rate at which a cognitive system, whether human, AI, or hybrid, moves its knowledge state $q(t)$ toward coherence with the generative manifold of truth $\tau(t)$:

$$\text{Alignment Velocity} = \frac{dq(t)}{dt} \rightarrow \tau(t)$$

This is not about how *much* content is produced, but how effectively a system discerns, refines, and internalizes meaningful insight.

6.1 $q(t) \rightarrow \tau(t)$: Tracking Knowledge State Toward Truth

In the Logostic model, each epistemic agent exists in a time-evolving state $q(t)$, which reflects their internal map of what is known. The target $\tau(t)$ represents the dynamically unfolding structure of reality—what is true, coherent, and generative at a given point in time.

To say that a scholar or a civilization is “advancing” is to say that $q(t)$ is converging toward $\tau(t)$ —that its map increasingly reflects the territory. But in a world of overproduction, high entropy, and noise, that convergence cannot be measured by publication volume alone.

Instead, we need metrics for how quickly insight reaches validity, how fast coherence stabilizes, and how meaning propagates across networks of belief.

Alignment velocity gives us a way to grade epistemic systems not by how much they say, but by how well they synchronize with truth.

6.2 Alignment Velocity vs Publication Volume

The publishing industry, bibliometrics, and academic prestige systems have historically been volume-centric:

- How many papers?
- How many citations?
- How many downloads, views, or altmetrics?

These are proxies for impact, but they are lagging and lossy.

By contrast, alignment velocity asks:

- How quickly does a valid idea emerge and become actionable?
- How often do new outputs reduce inconsistency or resolve prior ambiguity?
- Does a system improve its own epistemic coherence over time?

In this sense, ten deeply validated insights that collapse error surfaces in a field are more valuable than ten thousand plausible but unaligned outputs.

Without this reframing, civilizations may publish themselves into intellectual exhaustion while falling further behind in actual understanding.

6.3 How Fast Can a Civilization Converge on Valid Insight?

This is the ethical core of the Moore–Knowledge era. Not whether we *can* produce content, but whether we can align with truth fast enough to survive:

- Can scientific communities converge on correct models before crises emerge?
- Can AI systems learn to model truth faster than humans can deceive?
- Can hybrid epistemic agents elevate convergence rates without collapsing trust?

The velocity of alignment becomes the new cognitive survival metric for any knowledge ecosystem.

More provocatively, we might ask: is there an epistemic speed limit for civilization? A maximal dq/dt beyond which coherence destabilizes or ethical risk emerges? And if so, what *governs* it—hardware, cognition, social contract, or spiritual constraint?

These are no longer abstract questions. In the age of centaurs, exponential generation, and AI-assisted validation, the limiting factor is not speed of thought, but speed of convergence.

That convergence—between mind and model, between $q(t)$ and $\tau(t)$ —may ultimately define what it means for a civilization to be *intelligent* at scale.

7. Conclusion: Authorship Rewritten

We have entered a new epistemic era—one defined not by the scarcity of computational capacity, but by the scarcity of verified meaning. Moore’s Law has delivered us exponential compute; generative AI has transformed that capacity into near-limitless production. But the bottleneck has shifted decisively. It is no longer silicon—it is semantics.

In the age of centaur authorship, where human and machine co-generate and co-validate, the identity of the “author” is undergoing a profound transformation. Authorship is no longer reducible to the solitary voice, the named individual, or even the recognizable team. Instead:

The author is now a system—a recursive, distributed, epistemic process that spans human intuition, machine synthesis, algorithmic validation, and layered review.

This is not merely a change in writing style or citation practice. It is a redefinition of epistemic agency.

The hybrid author—the centaur—possesses emergent properties:

- Creativity informed by multi-domain correlation
- Validation enhanced by algorithmic speed and consistency
- Memory extended by retrieval-augmented generation
- Accountability shaped by human ethical constraints

In this system, the question is no longer “*Who wrote it?*” but “*What configuration of human–machine agency was responsible for validating it as true?*”

7.1 The Post-Moore Bottleneck is Meaning

The civilization shaped by Moore’s Law now faces a paradox: it can compute faster than it can comprehend. If knowledge is not just generated but *aligned*, then velocity alone is insufficient. Meaning must become the central engineering constraint in all epistemic systems.

Just as physical engineering must obey thermodynamic limits, knowledge engineering must obey semantic limits—the speed at which understanding, trust, and relevance can propagate through intelligent agents.

In this context, centaurs are not merely practical—they are structurally necessary. They are the only agents capable of operating across generation and validation, across speed and synthesis, without fracturing the human moral compass that grounds our definition of truth.

7.2 The Author is Now a System, Not a Self

As hybrid authorship becomes the norm, and as AI enters not just our keyboards but our review committees, our knowledge graphs, and our scientific reasoning pipelines, we must reckon with a new ontology of the author.

The author:

- May not be locatable in a body
- May span multiple modalities
- May evolve over time
- May include feedback loops of correction, dialogue, and reinforcement

What matters is not the *singularity* of the self, but the stability of epistemic function: does the system produce valid, trustworthy, convergent insight?

In this sense, authorship becomes a dynamic attractor, not a fixed point.

7.3 Toward Ethical Infrastructures for Epistemic Scale

Finally, if we accept that AI must participate in validation, and that authorship is hybrid, then we face an urgent ethical imperative: to build infrastructures of trust suited to this new scale.

These infrastructures must:

- Be transparent in how knowledge is generated and verified
- Preserve the traceability of human ethical oversight
- Incorporate uncertainty modeling, epistemic humility, and revision paths
- Recognize hybrid authorship in metadata, reputation systems, and credit models

We will need AI referees, human auditors, and layered responsibility protocols—not because AI is dangerous, but because epistemic velocity without ethical scaffolding invites collapse.

The ultimate question is not whether we trust AI. It is whether we can trust the centaur—this new system of thought, built from flesh and code, capable of knowledge at a scale that no previous civilization has approached.

We close with a reformulated maxim for this emerging age:

The future of knowledge belongs not to the fastest generator, but to the best-aligned system of convergence, coherence, and care.

That system is no longer human alone. It is human and AI—co-authors of our unfolding reality.

References

1. Moore, G. E. (1965). *Cramming more components onto integrated circuits*. *Electronics*, 38(8).
2. Bornmann, L., & Mutz, R. (2015). *Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references*. *Journal of the Association for Information Science and Technology*, 66(11), 2215–2222.
<https://doi.org/10.1002/asi.23329>
3. National Science Foundation. (2020). *Science & Engineering Indicators 2020*.
<https://ncses.nsf.gov/pubs/nsb20201>
4. OpenAI. (2023). *GPT-4 Technical Report*. <https://openai.com/research/gpt-4>
5. Korinek, A. (2023). *Language Models and Cognitive Automation for Economic Research*. *AEA Papers and Proceedings*, 113, 336–340.
6. Heaven, W. D. (2023). *AI is Dreaming Up Drugs That No One Has Ever Seen. Now We've Got to See if They Work*. MIT Technology Review. <https://www.technologyreview.com>
7. Yudkowsky, E. (2008). *Artificial Intelligence as a Positive and Negative Factor in Global Risk*. In Bostrom, N. & Ćirković, M. (Eds.), *Global Catastrophic Risks* (pp. 308–345). Oxford University Press.
8. Floridi, L., & Chiriatti, M. (2020). *GPT-3: Its Nature, Scope, Limits, and Consequences*. *Minds and Machines*, 30(4), 681–694. <https://doi.org/10.1007/s11023-020-09548-1>

The author has published over 4,500 papers at <https://www.researchgate.net/profile/Douglas-Youvan/research> . Google Scholar has indexed these preprints, and if you want a particular reference, the AI mode of a Google search (Gemini) has efficiently catalogued these preprints.